



# The relationship between study sponsorship, risks of bias, and research outcomes in atrazine exposure studies conducted in non-human animals: Systematic review and meta-analysis



L. Bero<sup>a</sup>, A. Anglemeyer<sup>b</sup>, H. Vesterinen<sup>c</sup>, D. Krauth<sup>d</sup>

<sup>a</sup> Charles Perkins Centre and Faculty of Pharmacy, University of Sydney, Australia

<sup>b</sup> Department of Collaborative Health and Human Services, California State University, Monterey Bay, USA

<sup>c</sup> University of California San Francisco, Program on Reproductive Health and the Environment, Oakland, USA

<sup>d</sup> Department of Clinical Pharmacy, University of California San Francisco, USA

## article info

### Article history:

Received 2 June 2015

Received in revised form 16 October 2015

Accepted 16 October 2015

Available online 13 December 2015

### Keywords:

Bias  
Conflicts of interest  
Chemical  
Funding source  
Atrazine  
Systematic review  
Meta-analysis  
Research synthesis  
Risk assessment  
Toxicology  
Animal research

## abstract

**Background:** A critical component of systematic review methodology is the assessment of the risks of bias of studies that are included in the review. There is controversy about whether funding source should be included in a risk of bias assessment of animal toxicology studies.

**Objective:** To determine whether industry research sponsorship is associated with methodological biases, the results, or conclusions of animal studies examining the effect of exposure to atrazine on reproductive or developmental outcomes.

**Methods:** We searched multiple electronic databases and the reference lists of relevant articles to identify original research studies examining the effect of any dose of atrazine exposure at any life stage on reproduction or development in non-human animals. We compared methodological risks of bias, the conclusions of the studies, the statistical significance of the findings, and the magnitude of effect estimates between industry sponsored and non-industry sponsored studies.

**Results:** Fifty-one studies met the inclusion criteria. There were no differences in methodological risks of bias in industry versus non-industry sponsored studies. 39 studies tested environmentally relevant concentrations of atrazine (11 industry sponsored, 24 non-industry sponsored, 4 with no funding disclosures). Non-industry sponsored studies (12/24, 50.0%) were more likely to conclude that atrazine was harmful compared to industry sponsored studies (2/11, 18.1%) ( $p$  value = 0.07). A higher proportion of non-industry sponsored studies reported statistically significant harmful effects (8/24, 33.3%) compared to industry-sponsored studies (1/11; 9.1%) ( $p$  value = 0.13). The association of industry sponsorship with decreased effect sizes for harm outcomes was inconclusive.

**Conclusion:** Our findings support the inclusion of research sponsorship as a risk of bias criterion in tools used to assess risks of bias in animal studies for systematic reviews. The reporting of other empirically based risk of bias criteria for animal studies, such as blinded outcome assessment, randomization, and all animals included in analyses, needs to improve to facilitate the assessment of studies for systematic reviews.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Results from animal studies are a critical, and often the only, input to assessing potential harm from exposure to chemicals. However, the lack of reproducibility of findings from animal research has reduced public confidence in the utility of animal experiments (van der Worp et al., 2010) and led to claims that animal research is a waste of financial resources (MacLeod et al., 2014). These problems with animal research have resulted in significant debate about how to assess biases in animal

studies used in systematic reviews, risk assessments and other regulatory decisions (Woodruff and Sutton, 2011; Rooney et al., 2014; National Academies of Science, 2014). A critical component of systematic review methodology is the assessment of the risks of bias of studies that are included in the review.

Risk of bias occurs when the methodological characteristics of a study produce a systematic error in the magnitude or direction of the results (Higgins and Green, 2011). Bias can shift effect estimates to be larger or smaller. For example, in controlled human clinical drug trials, studies with a high risk of bias (such as those lacking randomization, allocation concealment, or blinding of participants and outcome assessors) produce larger treatment effect sizes, thus falsely inflating the efficacy of the test interventions, compared to studies that have these

✉ Corresponding author at: D17, The Hub, 6th floor, Charles Perkins Centre, The University of Sydney, Camperdown, NSW 2006, Australia.  
E-mail address: [lisa.bero@sydney.edu.au](mailto:lisa.bero@sydney.edu.au) (L. Bero).

design features (Schulz and Grimes, 2002a; Schulz and Grimes, 2002b). However, biased human studies assessing the harms of drugs are more likely to report smaller estimates of adverse effects (Nieto et al., 2007).

Less is known about methodological risks of bias in animal studies, although a systematic review of instruments for assessing risks of bias in animal studies identified criteria that have been shown empirically to bias effect estimates in animal models (Krauth et al., 2013). For example, analyses of animal studies examining interventions for stroke, multiple sclerosis and trauma have shown that lack of randomization, blinding, specification of inclusion and exclusion criteria, statistical power, and failure to use comorbid animals are associated with inflated effect estimates of pharmaceutical interventions (Bebarta et al., 2003; Crossley et al., 2008; Sena et al., 2010a).

Industry funding for research and industry relationships with academic researchers pose an additional risk of bias. Considerable evidence shows a strong association between industry funding, investigator financial conflicts of interest, and biased outcomes in clinical research, even when controlling for methodological characteristics of the studies (Lundh et al., 2012). There is little evidence regarding the influence of these conflicts of interest on the outcomes of animal research (Krauth et al., 2014; Bennett et al., 2010; Abdel-Sattar et al., 2014). There are conflicting results concerning the association of industry funding and research outcomes among the small cohorts of animal studies that have been examined and further research on the influence of conflicts of interest on animal studies is needed (Bennett et al., 2010; Abdel-Sattar et al., 2014). There is controversy about whether funding source should be included in risk of bias assessments for studies included in systematic reviews (Bero, 2013).

Atrazine (6-chloro-N-ethyl-N'-(1-methylethyl)-1,3,5-triazine-2,4-diamine) is used as an herbicide. Atrazine is commonly found in drinking water in the United States. The EPA has concluded that "atrazine is an endocrine disruptor" (Agency, 2007) but not that atrazine affects amphibian sexual development (Agency, 2010). As of 2013, the EPA has not altered these conclusions (Agency, 2013). Atrazine studies are a good topic for an analysis of funding bias because concerns have been raised about the influence of industry sponsorship on the design and results of studies examining the effects of atrazine on reproductive and developmental outcomes (Hayes, 2004).

The objective of this study is to determine whether industry research sponsorship is associated with the methods, conclusions, or results of animal studies examining the effect of exposure to atrazine on reproductive or developmental outcomes. We test three specific hypotheses. First, we hypothesize that industry sponsored studies will be less likely to have conclusions indicating harm from atrazine than non-industry sponsored studies. Second, we test the hypothesis that industry sponsored studies will be less likely to report statistically significant results indicating harm from atrazine than non-industry sponsored studies. Third, we test the hypothesis that industry sponsored studies will have smaller effect estimates of harm than non-industry sponsored studies. In addition, we compare the methodological risks of bias of industry sponsored vs. non-industry sponsored studies to determine if there are differences in the methods of the studies.

## 2. Methods

We searched for studies that addressed the following question: "Does exposure to atrazine have adverse reproductive or developmental effects in non-human animals?" We searched for studies that had non-human animal subjects that were exposed to any dose of atrazine during any life stage. Exposure levels of atrazine were classified and adverse outcomes were grouped as described below.

### 2.1. Inclusion/exclusion criteria

Articles were included if they met the following criteria: (1) study conducted using whole animals; (2) original research, defined as a

study that presented original data and did not specifically state that it was a review; (3) atrazine compared to no exposure or control (eg, vehicle or some other exposure); (4) contains at least one group receiving only atrazine exposure; and (5) reports results data for at least one developmental and/or reproductive health outcome.

Studies were excluded if they met any of the following criteria: (1) pharmacokinetic or pharmacodynamic studies; (2) editorials, letters to the editor, commentaries, abstracts, unpublished reports, systematic reviews, meta-analyses; (3) studies comparing only different doses of atrazine; (4) studies in which atrazine was present in all the comparison groups; (5) in vitro analysis; (6) studies with no comparison groups.

Abstracts and article titles were first screened for inclusion. The full text of each article was then discussed by two authors who made a final decision about inclusion.

### 2.2. Search strategy

There were no language restrictions for the search. We searched Medline from January 1, 1966 to June 26th, 2013 using a search term combination containing the following MeSH terms, text words and word variants:

(atrazine) AND (animal\* OR preclinical OR "pre-clinical" OR mice OR rats OR rabbits OR dog OR dogs OR monkey OR monkeys OR "animal experimentation"[MeSH Terms] OR "models, animal"[MeSH Terms] OR "invertebrates"[MeSH Terms] OR "Animals"[MH] OR "animal population groups"[MeSH Terms]) NOT (humans[mh] NOT animals[mh: noexp]) AND (health effect OR health effects OR toxic OR toxicity OR toxicities OR efficacy OR efficacies OR toxicology OR safety OR harm\* OR drug effects[sh] OR therapeutic use[sh: noexp] OR adverse effects[sh] OR poisoning[sh] OR pharmacology[sh: noexp] OR chemically induced[sh]) AND eng[la] NOT review[pt] NOT systematic review\* NOT meta-analysis[pt].

We also searched, between May 1 and July 30, 2013 the following toxicology databases for articles that met our inclusion criteria:

DART <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?DARTETIC>  
EPA Science Inventory <http://www.epa.gov/gateway/science/>  
NIOSH TIC2 <http://www2.cdc.gov/nioshtic2/Nioshtic2.htm>  
Toxline <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE>  
USEPA Health and Environmental Studies Online <http://hero.epa.gov/>

TSCA Test Submissions: <http://www.ntis.gov/products/ots.aspx>

We identified 11 additional citations that were not in Medline. Nine could not be obtained even after contacting the authors. Two went on to full text screening. We searched the reference lists of all articles that met the inclusion criteria and identified one additional reference. Of the 3 additional references identified, one did not meet the inclusion criteria after full text screening.

### 2.3. Data extraction

#### 2.3.1. Single-coded data collection

DK collected the following characteristics from each included study:

#### 2.3.2. Study citation information

Title of the study, month of publication, year of publication, and journal name.

#### 2.3.3. Author affiliation

Author(s) affiliation(s) was obtained from the article and classified into (1) industry, if all authors were employed by industry (2) non-industry, if no author was employed by industry, or (3) combined, if at least one author was employed by industry and at least one author was not employed by industry. If a single author had affiliations with industry and non-industry sources, the study was coded as "combined".

### 2.3.4. Study design criteria

The following information was extracted from each study: (1) dose (s) of atrazine tested; (2) animal species and strain; (3) number of control and treated animals at the start of the study; (4) whether or not the investigator(s) assessed the effect of environmentally relevant concentrations (i.e. less than 1.89 ppm) of atrazine. We defined “environmentally relevant” as concentrations at or below 1.89 ppm. 18.9 ppb is considered the average concentration in community water systems according to the Agency for Toxic Substances and Disease Registry (ATSDR), and we multiplied that value by 100 to account for uncertainty and variability in the estimates.

### 2.3.5. Double-coded data collection

Two coders (DK and HV) independently collected data on characteristics of the studies that required judgment: methodological risks of bias, effect size and reported statistical significance of outcomes, study sponsorship source(s), investigator financial ties, and author conclusions. A third coder (LB) resolved any discrepancies.

## 2.4. Coding of study design criteria to assess risk of bias and other methodological criteria

We assessed risk of bias for the included studies using criteria derived from a review of tools for assessing bias in animal studies (Krauth et al., 2013).

For each publication, we coded each criterion as (1) yes, if the criteria were met (low risk of bias); (2) no, if the criteria were not met (high risk of bias); and when applicable (3) partial, if the criteria were partially met. The following criteria were assessed:

**Randomization** — Was the treatment randomly allocated to animal subjects so that each subject has an equal likelihood of receiving the intervention? Randomization was coded as (1) yes (2) no (3) partial.

**Concealment of allocation** — Were processes used to protect against selection bias by concealing from the investigators how treatment was allocated at the start of the study? Concealment of allocation was coded as (1) yes (2) no (3) partial.

**Blinded outcome assessment** — Was the investigator(s) involved with performing the experiment, collecting data, and assessing the outcome of the experiment unaware of which subjects received the treatment and which did not? Blinding was coded as (1) yes or (2) no.

**Statement of compliance with animal welfare requirements** — Did the author(s) state whether or not they complied with regulatory requirements for the handling and treatment of test animals? Statement of compliance with animal welfare requirements was coded as (1) yes or (2) no.

**Test animal description** — Did the author(s) describe in detail the test animal characteristics including, the animal species, strain, sub-strain, genetic background, age, supplier, sex, weight. At least one of these characteristics must be reported for this criterion to be coded as “yes.” Test animal description was coded as (1) yes or (2) no.

**Environmental parameters** — Did the author(s) adequately describe the housing and husbandry, nutrition, water, temperature, lighting conditions? At least one of these characteristics must be present for this criterion to be met. Environmental parameters were coded as (1) yes (2) no (3) partial.

**Sample size calculation** — Did the authors perform a sample size calculation to justify the total number of animals used in the study? Sample size calculation was coded as (1) yes or (2) no.

**Inclusion/exclusion criteria** — Were criteria used for including or excluding subjects specified? Inclusion/Exclusion criteria were coded as (1) yes (2) no (3) partial.

**Exposure levels justified** — Did the authors explain why the atrazine exposures levels were studied and whether a dose/response model was used to select the dose(s)? Exposure levels justified was coded as (1) yes or (2) no.

**Timing of exposure** (whether exposure period was sufficient to address outcome of interest) — Did the authors use a sufficient exposure period to address the outcome of interest? Timing of exposure was coded as (1) yes or (2) no.

**Optimal time window for outcome assessment** — Did the investigator provide sufficient time to pass before assessing the outcome? The optimal time window used in animal research should reflect the time needed to see the outcome. For example, it may take longer to see tumor development than molecular changes. Optimal time window for outcome assessment was coded as (1) yes (2) no (3) partial.

**All animals accounted for** — Did the investigator describe if animals were not included in the analysis and why they were not included? All animals accounted for was coded as (1) yes (2) no (3) partial.

### 2.4.1. Sponsorship source

Study sponsorship source as disclosed in each publication was coded as (1) any industry; (2) non-industry; (3) no sponsorship statement; and (4) no sponsorship.

### 2.4.2. Financial ties of authors

Investigator financial ties as disclosed in each publication were coded as (1) if at least one author of the study reported having a financial conflict of interest; (2) all authors reported having no conflicts of interest; (3) there was no disclosure statement. It was beyond the scope of this study to search for undisclosed author financial ties.

### 2.4.3. Role of the financial sponsor

For studies that disclosed a sponsor of any type, the role of the sponsor was coded as (1) not mentioned; (2) statement that the sponsor was not involved; (3) statement that the sponsor was involved. When applicable, we reported the sponsor's involvement (e.g., in the design, analysis or dissemination of the study).

## 2.5. Coding of conclusions of included studies

To test the hypothesis that industry sponsored studies would be less likely to have conclusions indicating harm from atrazine than non-industry sponsored studies, we coded the conclusions of each paper as reported in the abstract and conclusion sections. We coded author conclusions as (1) “harmful” if the overall conclusion suggested that atrazine was not safe at any of the doses tested; (2) “harmful at environmentally relevant concentrations” if the overall conclusion

suggested that atrazine was not safe at concentrations below 1.89 ppm (i.e. the upper limit of what is considered “environmentally relevant”); (3) “harmful at above environmentally relevant concentrations” if the author(s) only tested atrazine at high concentrations and concluded that atrazine was not safe at those high concentrations; (4) “neutral” if the paper did not draw a conclusion regarding the safety of atrazine or stated that the limitations of the study were so severe that the results were not valid; and (5) “no effect” if the overall conclusion suggested that atrazine was safe.

## 2.6. Coding of results of included studies

We report 1) the number of studies that reported statistically significant results indicating harm and 2) the magnitude of the harm effect estimates.

To test the hypothesis that industry sponsored studies would be less likely to report statistically significant results indicating harm from atrazine than non-industry sponsored studies, we assigned a code to each paper as follows.

### 2.6.1. Coding of individual outcome results

Only results for developmental or reproductive outcomes were recorded. Many different outcomes were reported and they were grouped into the following categories: 1) developmental abnormalities/mortality, 2) growth, 3) time to developmental landmark, 4) organ weight/size, 5) developmental enzymes, 6) reproductive hormones and 7) reproductive ability. See Supplemental File 1 for the specific outcomes included in each category.

For the purposes of this study, we coded a developmental or reproductive health related outcome result as “harmful” if the author(s) reported a statistically significant (for example,  $p < 0.05$ ) adverse reproductive or developmental health effect. We used the level of statistical significance reported in the paper. We coded that atrazine had “no effect” on reproductive or developmental health if the author(s) reported a statistically non-significant adverse effect (for example,  $p \geq 0.05$ ) or a statistically non-significant beneficial effect. We coded an effect as “beneficial” if atrazine showed a reproductive or developmental health benefit.

For studies that assessed a range of atrazine concentrations for a given outcome, we developed a decision rule for assigning a code for that outcome. If the only concentrations evaluated for an outcome were above environmentally relevant concentrations, we did not code the result as we are interested in the effect of atrazine at environmentally relevant concentrations. If an outcome was measured at concentrations both above and below 1.89 ppm, we coded only the values for the environmentally relevant concentrations. Therefore, if any of the exposure concentrations that tested harmful coincided with environmentally relevant concentrations, we coded the outcome as “harmful.”

### 2.6.2. Coding of results at the level of the study

A paper was coded as reporting “harmful” results if it reported more statistically significant outcomes demonstrating harmful effects than demonstrating no or not harmful effects; as “no effect” if it reported more outcomes that showed “no effect” compared to statistically significant harmful or not harmful outcomes; and as “neutral” if there an equal number of “no effect” and “harmful” outcomes.

## 2.7. Magnitude of effect estimates of included studies

We conducted meta-analyses to test the hypothesis that industry sponsored studies would have smaller effect estimates of harm than non-industry sponsored studies.

### 2.7.1. Data collection for meta-analysis

For each individual outcome, we collected the following data (often derived from tables, graphs, figures, etc.): measure of effect, confidence

interval, measure of variability (eg, standard deviation (SD) or standard error (SE)), p-value, statistical test used for each outcome, and the number of treated and untreated animals. If multiple time points were reported, we included all time points in the meta-analysis as to not assume a primary endpoint or arbitrarily assign an endpoint in the analysis.

### 2.7.2. Meta-analysis

We conducted a meta-analysis of the studies that had analyzable data. For a study to have analyzable data, it needed to report both a mean value and a measure of dispersion (standard error or standard deviation) or provide adequate data so that we could calculate these measures ourselves.

We calculated the effect of atrazine using a standardized mean difference (SMD) for each outcome. Due to the lack of independence of animals between outcomes within studies, we averaged SMDs and variances across outcomes for each study, yielding  $k$  average SMDs and variances for  $k$  studies. We pooled the data across studies and estimated summary average SMDs using random-effects models (DerSimonian and Laird, 1986). Specifically, we estimated the average SMD for each included study and used the inverse variance method to calculate study weights. The inverse variance method assumes that the variance for each study is inversely proportional to its importance; therefore, more weight is given to studies with less variance than studies with greater variance. The SMD null hypothesis ( $H_0$ : estimate = 0) states that there is no difference in effect of atrazine exposure on the specific outcomes when compared to a control.

For meta-analysis including outcomes for which the harmful direction is indicated as an increase, a number greater than zero suggests that atrazine is harmful when compared to the control exposure. Similarly, for meta-analyses including outcomes for which the harmful direction is indicated as a decrease, a number less than zero suggests that atrazine is harmful when compared to the control exposure.

We grouped outcomes within each outcome category according to whether an increase or decrease in the measure indicated a harmful outcome. A detailed description of whether a unit increase and decrease was considered to be harmful by outcome can be found in Supplemental Table 2. We stratified the direction of harm by increase and decrease because some outcome categories (e.g., reproductive hormones) are represented in both meta-analyses. For example, the meta-analyses of an outcome category (e.g., reproductive hormones) indicated by a unit increase causing harm (e.g. an increase in the hormone estradiol) and of a unit decrease causing harm (e.g. an increase in the hormone testosterone) contain independent data between meta-analyses.

We examined heterogeneity among the studies using the  $I^2$  statistic. We interpreted an  $I^2$  estimate greater than 50% as indicating moderate or high levels of heterogeneity. We anticipated high levels of heterogeneity as previous meta-analyses of animal studies have found high levels of heterogeneity between studies, potentially resulting from typical, small sample sizes in animal models (Macleod et al., 2004).

We evaluated differences in pooled effect estimates by sponsorship source to test the hypothesis that published industry sponsored studies are more likely to have outcomes that favor their sponsors (that is, outcome indicating less harm) than non-industry sponsored studies. We investigated the potential causes of heterogeneity by conducting a priori subgroup analyses using the  $\chi^2$  statistic with a significance level of 0.10. We conducted a subgroup analysis by type of outcome measure because we hypothesized that the effects of atrazine could vary by outcome and harm direction.

## 2.8. Additional statistical analysis

In addition to the effect estimates from the meta-analyses, we report the frequencies of each risk of bias criterion, study characteristics and the coding of the results and conclusions by sponsorship source.

### 3. Results

Our searches identified 51 studies that met the inclusion criteria (Fig. 1). There were 12 industry funded studies, 29 non-industry and 10 with no disclosures. The studies were published between 1984 and 2013 with the majority (41 studies) published between 2005 and 2010. Amphibians, fish and reptiles were most commonly studied ( $n = 39$ ), followed by rats ( $n = 8$ ), birds/fowl ( $n = 3$ ) and mice ( $n = 1$ ).

#### 3.1. Methodological risks of bias and study characteristics

There were no differences in methodological risks of bias in industry versus non-industry sponsored studies. The majority of studies gave a description of the test animals and the animal environment and reported that they were compliant with animal welfare requirements (Table 1). About half of the studies (26 of 51) reported that they were randomized, although none described concealment of allocation. Blinded outcome assessment, inclusion and exclusion criteria for the animals, sample size calculations, optimal time window investigated, and accounting for all animals were rarely or never reported. Industry sponsored studies were less likely to justify their choice of atrazine exposure levels (4/12 studies, 33%) compared to non-industry sponsored studies (19/29, 66%) ( $p$  value = 0.06).

Eighty percent (41/51) of studies had no disclosures about the financial conflicts of interest of the authors, although the majority of studies had authors who were not industry employees.

#### 3.2. Conclusions of studies

Of the 51 included studies, 39 (77%) reported that they tested environmentally relevant concentrations of atrazine: 11 of 12 (92%) industry sponsored studies, 24 of 29 (83%) non-industry sponsored studies, and 4 of the 10 (40%) studies with no funding disclosures.

Table 1  
Characteristics and risks of bias of included studies by sponsorship source ( $n = 51$ ).

Characteristic	Category	Sponsorship source		
		Any industry ( $N = 12$ )	Non-industry ( $N = 29$ )	No disclosure ( $N = 10$ )
Risk of bias <sup>3</sup>	Randomization	6	15	5
	Concealment of allocation	0	0	0
	Blinded outcome assessment	4	4	0
	Inclusion/exclusion criteria	0	2	0
	Sample size calculation	0	0	0
	Test animal description	12	29	10
	Animal environment described	12	29	9
	Exposure levels justified	4	19	5
	Optimal time window investigated	1	0	0
	All animals accounted for	2	3	2
	Compliant with animal welfare requirements	7	16	3
Financial conflict of interest	Yes	1	0	0
	No	4	4	1
	No disclosure	7	25	9
	Combined	3	1	1
Author affiliation	Non-industry employed	9	28	8
	Industry	0	0	1

As shown in Table 2, among studies that tested environmentally relevant concentrations of atrazine, non-industry sponsored studies (12/24, 50%) were more likely to conclude that atrazine was harmful compared to industry sponsored studies (2/11, 18%) ( $p$  value = 0.07) (Table 2).

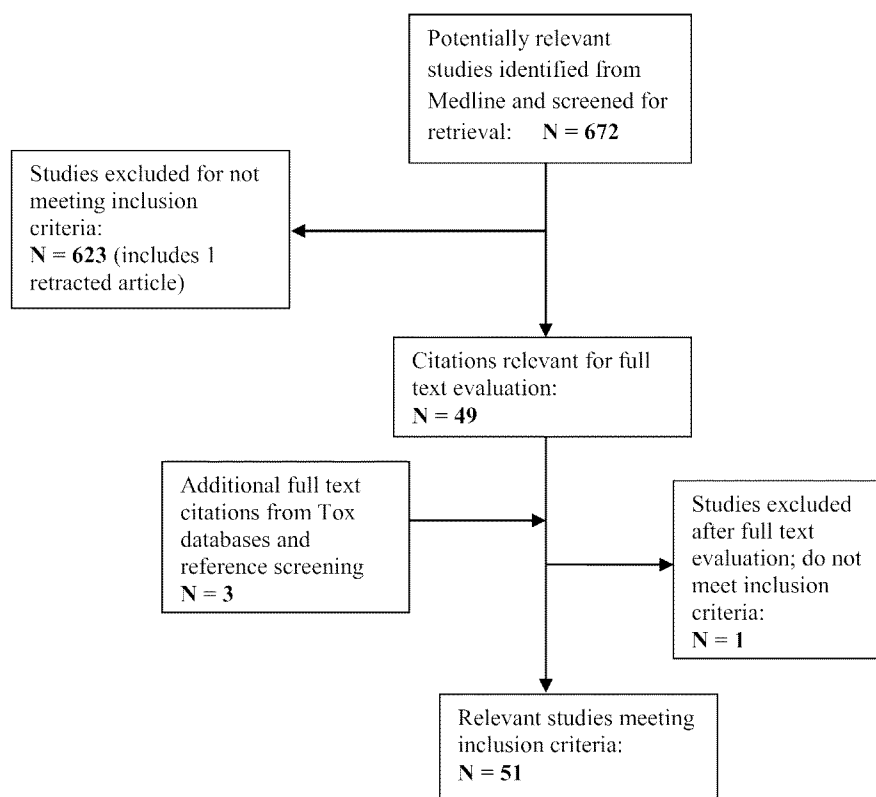


Fig. 1. Flow of included studies.

**Table 2**  
Results and conclusions of studies that tested environmentally relevant doses of atrazine by sponsorship source (n = 39).

	Direction of effect	Sponsorship source		
		Any industry (N = 11)	Non-industry (N = 24)	No disclosure (N = 4)
Results	Harmful	1	8	1
	No effect	8	13	3
	Neutral	2	3	0
Conclusions	Harmful	2	12	1
	No effect	8	10	2
	Neutral	1	2	1

### 3.3. Statistical significance of study results

As shown in Table 2, among studies that tested environmentally relevant concentrations of atrazine, a higher proportion of non-industry sponsored studies reported statistically significant harmful effects (8/24, 33%) compared to industry-sponsored studies (1/11; 9%) (p value = 0.13).

### 3.4. Magnitude of study results: meta-analyses of effect estimates

We identified 4 outcome categories for which a higher value indicated greater harm (Fig. 2). The outcome of developmental enzymes was not included in Fig. 2 as there was no variability in the sources of funding (n = 3). The pooled estimate for all studies examining outcomes where the direction of harm is indicated by an increase was 0.82 (95% CI 0.07, 1.57), showing a significant increase in risk of harm among atrazine exposed animals compared to control animals. Non-industry sponsored studies, in general, yielded stronger effects for outcomes in which the direction of harm was an increase than industry-sponsored studies (Fig. 2). For the outcome of time to developmental landmark, the pooled effect from industry-sponsored studies was significantly less than the pooled effect from non-industry sponsored studies, indicating less harm in the industry-sponsored studies (p value = 0.04).

We identified 5 outcome categories for which a lower value indicated harm (Fig. 3). The outcome of developmental enzymes was not included in Fig. 3 as there was insufficient variability in sponsorship. The pooled estimate for all studies examining the effects of atrazine on outcomes where the direction of harm is indicated by a decrease was

– 0.05 (95% CI – 0.49, 0.38) and yielded no significant effect. Industry-sponsored studies, in general, yielded stronger effects for outcomes in which the direction of harm was a decrease than non-industry sponsored studies.

## 4. Conclusions

We tested three specific hypotheses to determine whether industry sponsorship of research is associated with the results or conclusions of non-human animal studies examining the effect of atrazine exposure on reproductive or developmental outcomes. First, we tested the hypothesis that industry sponsored studies would be less likely to have conclusions indicating harm from atrazine than non-industry sponsored studies. This hypothesis was supported; 81% of industry supported studies did not conclude atrazine is harmful compared to 50% of non-industry supported studies.

Second, we tested the hypothesis that industry sponsored studies would be less likely to report statistically significant results indicating harm from atrazine than non-industry sponsored studies. This hypothesis was supported; 91% of industry sponsored studies did not report statistically significant outcomes for harm compared to 67% of non-industry sponsored studies.

However, the hypothesis that industry sponsored studies would have smaller effect estimates of harm than non-industry sponsored studies was not supported by the meta-analyses conducted.

Industry-sponsored studies, in general, yielded larger effects for outcomes in which the direction of harm was a decrease and smaller effects for outcomes in which the direction of harm was an increase than non-industry sponsored studies. Thus, the association of sponsorship with size of effect estimates is inconclusive. The exclusion of studies that did not assess atrazine at environmentally relevant concentrations and the small number of homogeneous outcome measures available for each meta-analysis are possible reasons that consistent differences in sizes of effect estimates between industry and non-industry sponsored studies were not observed.

We did not identify differences in methodological risks of bias between the industry and non-industry sponsored studies. The information needed to assess risks of bias was poorly reported. Blinded outcome assessment, inclusion and exclusion criteria for the animals, sample size calculations, optimal time window investigated, and accounting for all animals were rarely or never reported. About half of the studies reported that they were randomized, although none described concealment of allocation. Risk of bias criteria such as

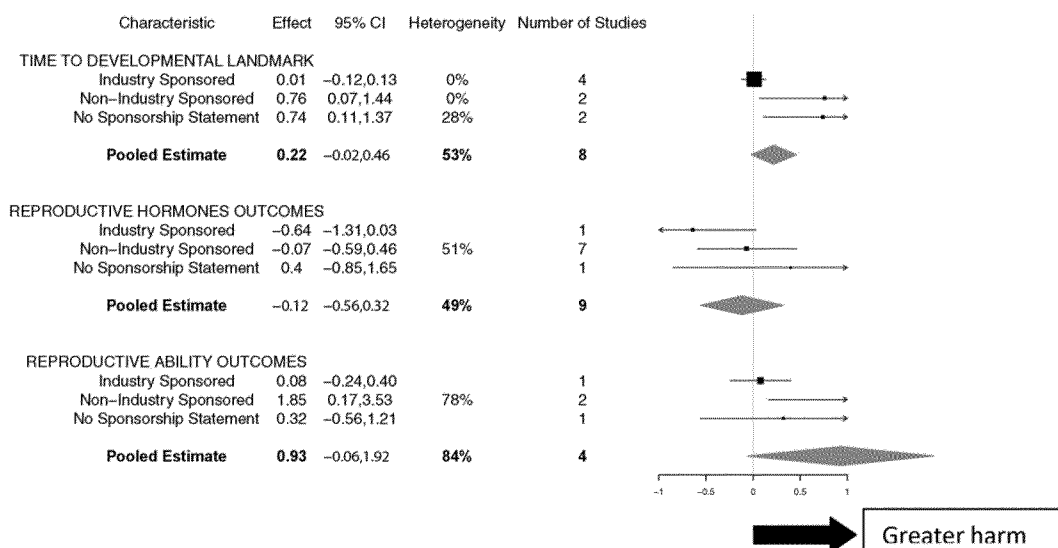


Fig. 2. Meta-analyses of the effect of atrazine on harm outcomes indicated by an increase in the effect size.

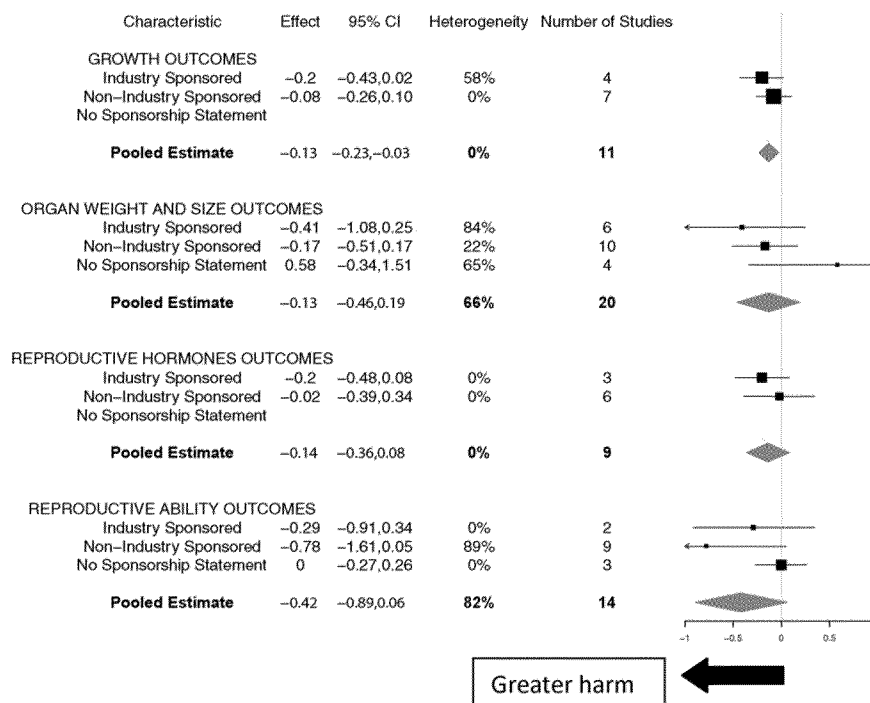


Fig. 3. Meta-analysis of the effect of atrazine on harm outcomes indicated by a decrease in effect size.

randomization and blinding should be consistently reported as there is empirical evidence that they affect the outcomes of animal research (Bebarta et al., 2003; Crossley et al., 2008; Sena et al., 2010b). In order to assess the quality of studies that will be included in systematic reviews all methodological risk of bias criteria must be fully reported. Furthermore, recent guidelines such as the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines recommend the reporting of selected risk of bias criteria (Kilkenny et al., 2010). Descriptive criteria, such as information on the test animals and the animal environment, were well reported in the included atrazine studies but are more relevant for determining compliance with regulatory requirements than assessing risks of bias (Krauth et al., 2013).

Since the differences in findings between industry and non-industry sponsored studies could not be explained by differences in the methodological risks of bias in the studies, another possible explanation is publication bias (MacLeod et al., 2004; Sena et al., 2010b; ter Riet et al., 2012) or selective outcome reporting. Publication bias is the failure to publish entire studies with statistically significant results and reporting bias is the failure to publish specific outcomes with statistically nonsignificant results. We could not assess the extent of selective outcome reporting in the studies included in our analysis because we had access only to the published reports and not the raw data.

Our findings suggest that reporting on conflicts of interest in animal toxicology studies also needs improvement. Eighty percent of the included articles failed to include any conflict of interest disclosures for the authors and twenty percent did not disclose the funding sources for the research. The reporting of funding sources is now common practice among clinical journals (Drazen et al., 2010). Given the prevalence of academic industry interactions in the field of toxicology and chemical risk assessment, similar requirements for reporting should be enforced for journals in environmental science and toxicology.

Complete and accurate information on exposure is critical for studies that will be used in systematic reviews assessing risks of chemical exposures. In this study, we classified exposures as environmentally relevant or not according to an accepted standard. We found that approximately three-quarters of the studies assessed atrazine at environmentally relevant exposures and limited our analyses to these studies. Although there were no differences in the proportions of industry sponsored

and non-industry sponsored studies that tested atrazine at environmentally relevant exposures, the industry sponsored studies were less likely to justify their choice of atrazine exposure levels.

Our study has a number of limitations. Although we searched multiple databases, we may have missed studies that were unpublished. The decision to exclude studies that did not test atrazine at environmentally relevant concentrations reduced our sample size. The heterogeneity of outcome measures in the included studies, even when grouped into similar categories, also limited the number of studies available for the meta-analyses testing the association of industry sponsorship with size of effect estimates.

In summary, industry sponsored studies were more likely to report findings that favored the sponsor. The non-industry sponsored studies were more likely to report statistically significant results and have conclusions indicating that atrazine is harmful compared to industry sponsored studies. This finding supports the inclusion of research sponsorship as a risk of bias criteria in tools used to assess risks of bias in animal studies for systematic reviews. The reporting of other risk of bias criteria for animal studies, such as blinded outcome assessment, randomization, and all animals included in analyses, needs to improve to facilitate the assessment of studies for systematic reviews.

#### Funding source

National Institute of Environmental Health Sciences (Grant # R21ES021028) (<http://www.niehs.nih.gov>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### Conflict of interest

The authors declare no competing conflicts of interest.

#### Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.envint.2015.10.011>

## References

- Abdel-Sattar, M., et al., 2014. The relationship between risk of bias criteria, research outcomes, and study sponsorship in a cohort of preclinical thiazolidinedione animal studies: a meta-analysis. *Evid Based Preclin Med* 1 (1), 11–20.
- Agency, E.P., Atrazine chemical summary. 2007. p. 12.
- Agency, E.P. Atrazine updates: amphibians. 2010 July 23, 2014]; Available from: [http://www.epa.gov/pesticides/reregistration/atrazine/atrazine\\_update.htm#amphibian](http://www.epa.gov/pesticides/reregistration/atrazine/atrazine_update.htm#amphibian)
- Agency, E.P. Atrazine updates: atrazine evaluation process. 2013 July 23, 2014]; Available from: [http://www.epa.gov/pesticides/reregistration/atrazine/atrazine\\_update.htm#atrazine](http://www.epa.gov/pesticides/reregistration/atrazine/atrazine_update.htm#atrazine)
- Bebarta, V., Luyten, D., Heard, K., 2003. Emergency medicine animal research: does use of randomization and blinding affect the results? *Acad. Emerg. Med.* 10 (6), 684–687.
- Bennett, C.L., et al., 2010. Association between pharmaceutical support and basic science research on erythropoiesis-stimulating agents. *Arch. Intern. Med.* 170 (16), 1490–1498.
- Bero, L.A., 2013. Why the Cochrane risk of bias tool should include funding source as a standard item. *Cochrane Database Syst. Rev.* 12, ED000075.
- Crossley, N.A., et al., 2008. Empirical evidence of bias in the design of experimental stroke studies: a metaepidemiologic approach. *Stroke* 39 (3), 929–934.
- DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. *Control. Clin. Trials* 7 (3), 177–188.
- Drazen, J.M., et al., 2010. Toward more uniform conflict disclosures: the updated ICMJE conflict of interest reporting form. *Natl. Med. J. India* 23 (4), 196–197.
- Hayes, T., 2004. There is no denying this: Defusing the confusion about Atrazine. *Bioscience* 54 (12), 1138–1149.
- Higgins, J., Green, S., 2011. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0. John Wiley and Sons, Chichester, UK.
- Kilkenny, C., et al., 2010. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol.* 8 (6), e1000412.
- Krauth, D., Woodruff, T.J., Bero, L., 2013. Instruments for assessing risk of bias and other methodological criteria of published animal studies: a systematic review. *Environ. Health Perspect.* 121 (9), 985–992.
- Krauth, D., et al., 2014. Nonindustry-sponsored preclinical studies on statins yield greater efficacy estimates than industry-sponsored studies: a meta-analysis. *PLoS Biol.* 12 (1), e1001770.
- Lundh, A., et al., 2012. Industry sponsorship and research outcome. *Cochrane Database Syst. Rev.* 12, MR000033.
- Macleod, M.R., et al., 2004. Use of animal experimental data reveals influence of study design and publication bias. *Stroke* 35 (5), 1203–1208.
- Macleod, M.R., et al., 2014. Biomedical research: increasing value, reducing waste. *Lancet* 383 (9912), 101–104.
- National Academies of Science. Review of EPA's Integrated Risk Information System (IRIS) Process. 2014, Washington (DC).
- Nieto, A., et al., 2007. Adverse effects of inhaled corticosteroids in funded and nonfunded studies. *Arch. Intern. Med.* 167 (19), 2047–2053.
- Rooney, A.A., et al., 2014. Systematic review and evidence integration for literature-based environmental health science assessments. *Environ. Health Perspect.* 122 (7), 711–718.
- Schulz, K.F., Grimes, D.A., 2002a. Blinding in randomised trials: hiding who got what. *Lancet* 359 (9307), 696–700.
- Schulz, K.F., Grimes, D.A., 2002b. Allocation concealment in randomised trials: defending against deciphering. *Lancet* 359 (9306), 614–618.
- Sena, E.S., et al., 2010a. Factors affecting the apparent efficacy and safety of tissue plasminogen activator in thrombotic occlusion models of stroke: systematic review and meta-analysis. *J. Cereb. Blood Flow Metab.* 30 (12), 1905–1913.
- Sena, E.S., et al., 2010b. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS Biol.* 8 (3), e1000344.
- ter Riet, G., et al., 2012. Publication bias in laboratory animal research: a survey on magnitude, drivers, consequences and potential solutions. *PLoS One* 7 (9), e43404.
- van der Worp, H.B., et al., 2010. Can animal models of disease reliably inform human studies? *PLoS Med.* 7 (3), e1000245.
- Woodruff, T.J., Sutton, P., G. Navigation Guide Work, 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health Aff (Millwood)* 30 (5), 931–937.



Supplemental Table 1: Categorization of Outcomes measured in included Atrazine studies

<b>1. Developmental Abnormalities/Mortality</b>
Abnormal development in the offspring
Gonad abnormality
Left -to-right testis ratio
<b>2. Growth</b>
Fat body size
Fish length
Body length
Growth length
Length
Smolt length
Smolt weight
Snout to vent length
Terminal body length
Mass of f2 larvae
Mass of versicolor tadpoles
3 month growth rate in offspring
Birth weight
Larvae wet weight
ash free dry weight
larval period (days)
Percent survival metamorphosis
Proportion of tadpoles initiate metamorphosis

<b>3. Time to Developmental Landmark</b>
Number of days between cessation of treatment and first prooesterus
Time to metamorphosis
age at metamorphosis
age at preputial separation
days to reach stage 66 of first larva
duration of time between maturity achievement
time to reach stage 42
time to TR
<b>4. Organ weight / size</b>
absolute testis weight
Gonad surface area
anogenital index
Diameter of immature follicles
Diameter of mature follicles
Diameter: testis weight (ratio)
Laryngeal muscle area
Ovary weight
Oviduct weight
Prostate weight
Ratio of testis to body weight
Relative adrenal weight
Seminal vesicle weight
Uterus weight

Ventral prostrate weight
Vesicle weight
Hepatosomatic index
Gonadal somatic index
Breeding gland area
<b>5. Developmental enzymes</b>
aromatase activity
gonad aromatase
mRNA of aromatase in brain
mRNA of luteinizing hormone in brain
mRNA of follicle stimulating hormone
<b>6. Reproductive hormones</b>
Circulating concentrations of estradiol
Circulating concentrations of luteinizing hormone (LH)
Circulating concentrations of testosterone
Estradiol concentration
Intratesticular testosterone concentration
Plasma aldosterone
Plasma corticosterone
Plasma E2
Plasma estradiol
Plasma LH
Plasma testosterone
Serum androstenedione

Serum estradiol
Serum estrone
Serum prolactin
Serum testosterone
Testicular estradiol
Testicular testosterone
<b>7. Reproductive ability</b>
Fecundity
Fertility
Fertilization rate
Hatch success
Locomotion activity
Mass corrected number of eggs released per female
Molting period
Number of mating attempts
Abnormal sperm rate
Daily sperm rate
Daily sperm/testes
Sperm motility
Sperm / area
Number of eggs
Number of embryos
Number of offspring
Number of spawns

Number of testicular ovarian follicles
Display of courtship
Number of implantations
Percent adult emergence
Percent hatch of eggs
Percent postimplant loss
Percent preimplant loss
Rate of travel
Resting spermatogonia
Seminiferous tubule diameter
Sex ratio
Spermatids
Spermatogonia
Percent survival in different frog species
Swimming speed
Total offspring
Dead sperm
Epidymal sperm number
Number of corpora lutea

Supplemental Table 2: Outcome Groups and Direction of Harm

HARMFUL DIRECTION- INCREASE	HARMFUL DIRECTION- DECREASE
<b>Developmental Abnormalities/Mortality</b>	
Abnormal development – in the offspring	-
Gonad abnormality	-
<b>Time to Developmental Landmark</b>	
Number of days between cessation of tx and first prooosterus	-
Time to metamorphosis	-
age at metamorph	-
age at preputial separation	-
days to reach stage 66 of first larva	-
duration of time between maturity achievement	-
time to reach stage 42	-
time to TR	-
<b>Organ weight / size</b>	
Diameter of immature follicles	absolute testis weight
Diameter of mature follicles	Gonad surface area
Relative adrenal weight	anogenital index
Hepatosomatic index	Diameter: testis weight (ratio)
-	Laryngeal muscle area

-	Ovary weight
-	Oviduct weight
-	Prostate weight
-	Ratio of testis to body weight
-	Seminal vesicle weight
-	Uterus weight
-	Ventral prostate weight
-	Vesicle weight
-	Gonadal somatic index
-	Breeding gland area
<b>Reproductive hormones</b>	
Circulating concentrations of estradiol	Circulating concentrations of luteinizing hormone (LH)
Estradiol concentration	Circulating concentrations of testosterone
plasma corticosterone	Intratesticular testosterone concentration
plasma E2	plasma aldosterone
plasma estradiol	plasma LH
serum estradiol	plasma testosterone
serum estrone	serum testosterone
testicular estradiol	testicular testosterone
<b>Reproductive ability</b>	
Molting period	Fecundity
Abnormal sperm rate	Fertility

Number of testicular ovarian follicles	Fertilization rate
Percent postimplant loss	Hatch success
Percent preimplant loss	Locomotion activity
Dead sperm	Mass corrected number of eggs released per female
-	Number of mating attempts
-	Daily sperm rate
-	Daily sperm/testes
-	Sperm motility
-	Sperm / area
-	Number of eggs
-	Number of embryos
-	Number of offspring
-	Number of spawns
-	Display of courtship
-	Number of implantations
-	Percent adult emergence
-	Percent hatch of eggs
-	Rate of travel
-	Resting spermatogonia
-	Seminiferous tubule diameter
-	Spermatids
-	Spermatogonia
-	Swimming speed



-	Total offspring
-	Epidymal sperm number
<b>Growth</b>	
-	Fat body size
-	Fish length
-	Body length
-	Growth length
-	Length
-	Smolt length
-	Smolt weight
-	Snout to vent length
-	Terminal body length
-	Mass of f2 larvae
-	Mass of versicolor tadpoles
-	3 month growth rate – in offspring
-	Birth weight
-	Larvae wet weight
-	ash free dry weight
-	larval period (days)
-	Percent survival metamorph
-	Proportion of tadpoles initiate metamorph
<b>Developmental enzymes</b>	
-	aromatase activity

-	gonad aromatase
-	mRNA of aromatase in brain
-	mRNA of luteinizing hormone in brain
-	mRNA of follicle stimulating hormone